

RAGFlow: Advanced Document Processing System

RAGFlow is a cutting-edge document processing system that combines the power of vector databases with large language models to create an intelligent document search and question-answering platform.

Key Features:

- Vector-based document search using OpenAI embeddings
- Intelligent text chunking with overlap for better context
- Deduplication to prevent redundant content
- Caching system for faster processing
- MMR (Maximum Marginal Relevance) reranking for diverse results
- Support for PDF, TXT, and Markdown files
- Real-time chat interface for document queries

Technical Architecture:

The system uses ChromaDB as the vector database, OpenAI's text-embedding-3-small model for generating embeddings, and Claude for generating human-like responses. Documents are automatically chunked into manageable pieces with configurable overlap to maintain context across boundaries.

Use Cases:

- Research document analysis
- Knowledge base management
- Customer support automation
- Educational content processing
- Legal document review
- Technical documentation search

The system is designed to handle large documents efficiently while maintaining high accuracy in retrieval and generation tasks.